Randomly evolving trees II

L. Pál,

KFKI Atomic Energy Research Institute 1525 Budapest, P.O.B. 49. Hungary

November 5, 2002

Abstract

Generating function equation has been derived for the probability distribution of the number of nodes with $k \geq 0$ outgoing lines in randomly evolving special trees defined in an earlier paper arXiv:cond-mat/0205650. The stochastic properties of the end-nodes ($k = 0$) have been analyzed, and it was shown that the relative variance of the number of end-nodes vs. time has a maximum when the evolution is either subcritical or supercritical. On the contrary, the time dependence of the relative dispersion of the number of dead end-nodes shows a minimum at the beginning of the evolution independently of its type. For the sake of better understanding of the evolution dynamics the survival probability of random trees has been investigated, and asymptotic expressions have been derived for this probability in the cases of subcritical, critical and supercritical evolutions. In critical evolution it was shown that the probability to find the tree lifetime larger than x, is decreasing to zero as $1/x$, if $x \to \infty$. Approaching the critical state it has been found the fluctuations of the tree lifetime to become extremely large, and so near the critical state the average lifetime could be hardly used for the characterization of the process.

PACS: 02.50.-r, 02.50.Ey, 05.40.-a

1 Introduction

In a previous paper [1] we defined and analyzed random processes with continuous time parameter describing the evolution of special trees consisting of living and dead nodes connected by lines. The initial state $\mathcal{S}_0$ of the tree

corresponds to a single living node called root which at the end of its life capable to produce new living nodes, and after that it becomes immediately dead. The new nodes are promptly connected to the dead node and each of them *independently of the others* can evolve further like a root. It is evident that the random evolution of this type is nothing else than a branching process.

In what follows we will use the notations applied in [1]. Therefore, the distribution function of the lifetime τ of a living node will be denoted by $T(t)$, and the probability that the number ν of living nodes produced by one dying precursor is equal to j by f_j where $j \in \mathcal{Z}$. [1]

In order to characterize the tree evolution two non-negative integer valued random functions $\mu_\ell(t)$ and $\mu_d(t)$ were introduced in [1]. $\mu_\ell(t)$ is the number of living nodes, while $\mu_d(t)$ is that of dead nodes at $t \geq 0$. It was mentioned also that the trivial equality

$$\mu_\ell(t) + \mu_d(t) = \mu_e(t) + 1$$

must be valid with probability 1 at any $t \geq 0$. Here $\mu_e(t)$ is the number of lines in the tree at the time moment $t \geq 0$. For the generating functions

$$g^{(\ell)}(t,z) = \sum_{n=0}^{\infty} p^{(\ell)}(t,n)\, z^n, \quad \text{és} \quad g^{(d)}(t,z) = \sum_{n=0}^{\infty} p^{(d)}(t,n)\, z^n, \tag{1}$$

where

$$p^{(\ell)}(t,n) = \mathcal{P}\{\mu_\ell(t) = n|\mathcal{S}_0\} \quad \text{és} \quad p^{(d)}(t,n) = \mathcal{P}\{\mu_d(t) = n|\mathcal{S}_0\}$$

the following integral equations were derived:

$$g^{(\ell)}(t,z) = [1 - T(t)]\, z + \int_0^t q\left[g^{(\ell)}(t-t',z)\right]\, dT(t') \tag{2}$$

and

$$g^{(d)}(t,z) = 1 - T(t) + z \int_0^t q\left[g^{(d)}(t-t',z)\right]\, dT(t'), \tag{3}$$

where

$$q(z) = \sum_{j=0}^{\infty} f_j\, z^j.$$

[1] $\mathcal{Z}$ is the set of non-negative integers.

If $T(t) = 1 - e^{-Qt}$, then the evolution becomes a Markov process.

The nodes can be sorted into groups according to the number of outgoing lines. Denote by $\mu(t,k)$ the number of nodes with $k \geq 0$ outgoing lines at the time instant $t \geq 0$. A node not having outgoing line is called *end-node*. It is obvious that an end-node could be either live or dead. Therefore, the number of end-nodes $\mu(t,0)$ can be written as a sum of numbers of living and dead end-nodes, i.e.

$$\mu(t,0) = \mu_\ell(t,0) + \mu_d(t,0).$$

Since all living nodes are end-nodes $\mu_\ell(t,0)$ can be replaced by $\mu_\ell(t)$. The total number of dead nodes $\mu_d(t)$ is given by

$$\mu_d(t) = \sum_{k=0}^{\infty} \mu_d(t,k).$$

In what follows we will calculate the probability distribution of $\mu_d(t,k)$ and investigate the properties of end-nodes $\mu(t,0)$ playing important role in random tree evolution.

In order to have a deeper insight into the dynamics of the evolution process, we will derive an important equation determining the probability distribution function of the *tree lifetime*.

2 Generating functions

2.1 Distribution of ν

The basic properties of the probability distribution of the number of nodes in a random tree are depending mainly on the distribution law of the number ν of living nodes produced by one dying precursor. In the sequel we will use the notations

$$\mathbf{E}\{\nu\} = q_1 \qquad \text{and} \qquad \mathbf{D}^2\{\nu\} = q_2 + q_1 - q_1^2$$

introduced already in [1] for the expectation and the variance of ν where

$$q_j = \left[\frac{d^j q(z)}{dz^j}\right]_{z=1}, \qquad j = 1, 2, \ldots$$

are the factorial moments of ν. It was shown in [1] that the time dependence of the random evolution is determined almost completely by the expectation value q_1. The evolution is called subcritical if $q_1 < 1$, critical if $q_1 = 1$ and supercritical if $q_1 > 1$. In the further considerations we are going to use four simple distributions for the random variable ν.

2.1.1 Arbitrary distribution

It has been shown in [1] that the equations derived for the first and the second moments of $\mu_\ell(t), \mu_d(t), \ldots$ are true for any distribution of ν provided that the moments q_1 and q_2 are finite. This type of distributions of ν is called *arbitrary* and will be denoted by $\mathbf{a}$.

Many times it is expedient to assume distributions of ν to be completely determined by one or two parameters.

2.1.2 Geometric and Poisson distributions

As known the geometric and Poisson distributions are containing one parameter only, and so we have

$$\mathcal{P}\{\nu = j\} = \begin{cases} \frac{1}{1+q_1} \left(\frac{q_1}{1+q_1}\right)^j, & \text{if } \nu \in \mathbf{g}, \\ \\ e^{-q_1} \frac{q_1^j}{j!}, & \text{if } \nu \in \mathbf{p}, \end{cases}$$

where $\mathbf{g}$ and $\mathbf{p}$ refer to the geometric and the Poisson distributions, respectively. For the sake of completeness we write

$$q(z) = \begin{cases} \frac{1}{1+(1-z)q_1}, & \text{if } \nu \in \mathbf{g}, \\ \\ e^{-(1-z)q_1}, & \text{if } \nu \in \mathbf{p}, \end{cases}$$

and

$$\mathbf{E}\{\nu\} = \begin{cases} q_1, & \text{if } \nu \in \mathbf{g}, \\ \\ q_1, & \text{if } \nu \in \mathbf{p}, \end{cases} \qquad \text{while} \qquad \mathbf{D}^2\{\nu\} = \begin{cases} q_1(1+q_1), & \text{if } \nu \in \mathbf{g}, \\ \\ q_1, & \text{if } \nu \in \mathbf{p}. \end{cases}$$

2.1.3 Truncated arbitrary distribution

In this case the possible values of the random variable ν are $0, 1$ and 2 with probabilities f_0, f_1 and f_2, respectively. This distribution will be denoted by $\mathbf{t}$. The corresponding generating function of ν is given by

$$q(z) = f_0 + f_1 z + f_2 z^2 = 1 + q_1(z-1) + \frac{1}{2} q_2 (z-1)^2.$$

This choice of $q(z)$ has a great advantage, it makes possible to obtain exact solutions of the generating function equations of $\mu_\ell(t), \mu_d(t), \ldots$ characterizing the tree evolution.

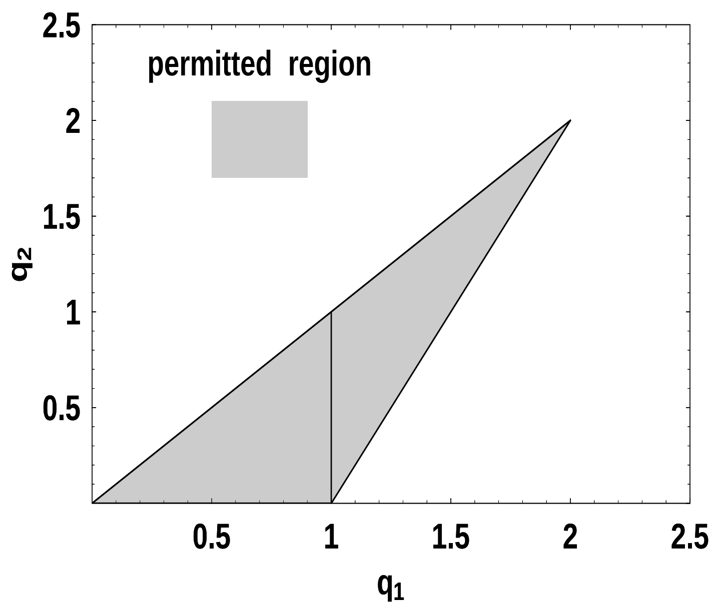


Figure 1: The permitted values of q_1 and q_2 in the case of distribution $\mathbf{t}$ of ν.

It seems to be useful to cite the following trivial relations:

$$f_0 = 1 - q_1 + \frac{1}{2}\, q_2, \quad f_1 = q_1 - q_2, \quad f_2 = \frac{1}{2}\, q_2,$$

which follow from the equations

$$f_0 + f_1 + f_2 = 1, \quad f_1 + 2f_2 = q_1, \quad 2f_2 = q_2,$$

where $q_2 = \mathbf{D}^2\{\nu\} - q_1 + q_1^2$. Since f_0, f_1 and f_2 are non-negative, smaller than 1, real numbers and their sum is equal to 1, the possible values of q_1 and q_2 are restricted. The permitted values of q_1 and q_2 are shown in Fig. 1. (See the shaded triangle!)

2.2 Distribution of the number of dead nodes with k outgoing lines

As has been already mentioned, a living node may create $k \geq 0$ new living nodes and after that it becomes immediately dead. This node is called *dead node of out-degree* k. It was introduced the random function $\mu_d(t,k)$ giving the number of these dead nodes at time instant $t \geq 0$. Now, we want to determine the probability that $\mu_d(t,k)$ is equal to $n \geq 0$ provided that at $t = 0$ the tree was in the state $\mathcal{S}_0$. First of all, we define the probability generating function

$$g_k^{(d)}(t,z) = \sum_{n=0}^{\infty} p_k^{(d)}(t,n)\ z^n, \tag{4}$$

where

$$p_k^{(d)}(t,n) = \mathcal{P}\{\mu_d(t,k) = n|\mathcal{S}_0\}$$

is the probability that at time $t \geq 0$ the tree has exactly n dead nodes with $k \geq 0$ outgoing lines provided that at $t = 0$ it was in its initial state $\mathcal{S}_0$. By using similar considerations as we did in [1], we have

$$p_k^{(d)}(t,n) = e^{-Qt}\ \delta_{n,0} +$$

$$Q \int_0^{\infty} e^{-Q(t-t')} \left\{ f_0 \left[\delta_{0,k}\delta_{n,1} + (1-\delta_{0,k})\delta_{n,0}\right] + \sum_{j=1}^{\infty} f_j R_{j,k}^{(d)}(t',n) \right\}\ dt', \tag{5}$$

where

$$R_{j,k}^{(d)}(t',n) =$$

$$\delta_{j,k} \sum_{n_1+\cdots+n_j=n-1} \prod_{i=1}^{j} p_k^{(d)}(t',n_i) + (1-\delta_{j,k}) \sum_{n_1+\cdots+n_j=n} \prod_{i=1}^{j} p_k^{(d)}(t',n_i).$$

The expression in square brackets at f_0 reflects that two mutually excluding possibilities exist depending on whether $k = 0$ or $k \neq 0$. One can see immediately that the generating function defined by Eq. (4) satisfies the equation

$$g_k^{(d)}(t,z) = e^{-Qt} + Q \int_0^t e^{-Q(t-t')} f_0 [1 - (1-z)\ \delta_{k,0}]\ dt' +$$

$$Q \int_0^t e^{-Q(t-t')} \left\{ \sum_{j=1}^{\infty} f_j (1 - \delta_{j,k}) \left[g_k^{(d)}(t',z) \right]^j + \sum_{j=1}^{\infty} f_j \delta_{j,k} z \left[g_k^{(d)}(t',z) \right]^j \right\} dt'.$$

By rearranging the right side we find

$$g_k^{(d)}(t,z) = e^{-Qt} - (1-z) f_0\ \delta_{k,0}\ Q \int_0^t e^{-Q(t-t')}\ dt' -$$

$$Q \int_0^t e^{-Q(t-t')} \left\{ f_k\ (1-z)\ (1-\delta_{0,k})\ \left[g_k^{(d)}(t',z) \right]^k - q \left[g_k^{(d)}(t',z) \right] \right\} dt'. \quad (6)$$

When $k > 0$ then the equation (6) takes the form:

$$g_k^{(d)}(t,z) = e^{-Qt} +$$

$$Q \int_0^t e^{-Q(t-t')} \left\{ -f_k (1-z)\ \left[g_k^{(d)}(t',z) \right]^k + q \left[g_k^{(d)}(t',z) \right] \right\} dt'. \quad (7)$$

The differential equation equivalent to (6) is nothing else than

$$\frac{\partial g_k^{(d)}(t,z)}{\partial t} = -Q\ g_k^{(d)}(t,z) - Q\ f_0 (1-z)\ \delta_{k,0} -$$

$$Q \left\{ f_k\ (1-z)\ (1-\delta_{0,k})\ \left[g_k^{(d)}(t,z) \right]^k - q \left[g_k^{(d)}(t,z) \right] \right\} \quad (8)$$

and the initial condition is given by $\lim_{t \downarrow 0} g_k^{(d)}(t,z) = 1$.

2.3 Joint distribution of the numbers of two dead nodes with different out-degrees

The following step in our considerations is the determination of the joint probability distribution

$$\mathcal{P}\{\mu_d(t,k_1) = n_1, \mu_d(t,k_2) = n_2|\mathcal{S}_0\} = p^{(d)}_{k_1,k_2}(t,n_1,n_2), \tag{9}$$

where $k_1 \neq k_2$. It is clear from the definition (9) that $p^{(d)}_{k_1,k_2}(t,n_1,n_2)$ is the probability that in the time interval $(0,t)$ the evolution produces n_1 nodes with k_1 and n_2 nodes with k_2 outgoing lines provided that at the moment $t=0$ the tree was in its initial state $\mathcal{S}_0$. By using similar arguments applied in deriving the backward equation (6), we obtain for the generating function

$$g^{(d)}_{k_1,k_2}(t,z_1,z_2) = \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} p^{(d)}_{k_1,k_2}(t,n_1,n_2)\; z_1^{n_1}\; z_2^{n_2} \tag{10}$$

the following equation when $k_1 \neq k_2$:

$$g^{(d)}_{k_1,k_2}(t,z_1,z_2) = e^{-Qt} - Q\, f_0 \int_0^t e^{-Q(t-t')}[(1-z_1)\,\delta_{k_1,0} + (1-z_2)\,\delta_{k_2,0}]\, dt' -$$

$$Q \int_0^t e^{-Q(t-t')}(1-z_1)\, f_{k_1}\, \left[g^{(d)}_{k_1,k_2}(t',z_1,z_2)\right]^{k_1}\, dt' -$$

$$Q \int_0^t e^{-Q(t-t')}(1-z_2)\, f_{k_2}\, \left[g^{(d)}_{k_1,k_2}(t',z_1,z_2)\right]^{k_2}\, dt' +$$

$$Q \int_0^t e^{-Q(t-t')} q[g^{(d)}_{k_1,k_2}(t',z_1,z_2)]\, dt'. \tag{11}$$

It follows from this equation that

$$\lim_{z_2\to 0} g^{(d)}_{k_1,k_2}(t,z_1,z_2) = g^{(d)}_{k_1}(t,z_1) \qquad \text{and} \qquad \lim_{z_1\to 0} g^{(d)}_{k_1,k_2}(t,z_1,z_2) = g^{(d)}_{k_2}(t,z_2).$$

For the sake of the completeness we are giving here the differential equation equivalent to (11). It has the form:

$$\frac{\partial g^{(d)}_{k_1,k_2}(t,z_1,z_2)}{\partial t} = -Q\; g^{(d)}_{k_1,k_2}(t,z_1,z_2) + Q\; q[g^{(d)}_{k_1,k_2}(t,z_1,z_2)] -$$

$$Q\ f_0[(1-z_1)\ \delta_{k_1,0} + (1-z_2)\ \delta_{k_2,0}] -$$

$$(1-z_1)\ f_{k_1}\ \left[g^{(d)}_{k_1,k_2}(t,z_1,z_2)\right]^{k_1} - (1-z_2)\ f_{k_2}\ \left[g^{(d)}_{k_1,k_2}(t,z_1,z_2)\right]^{k_2}, \qquad (12)$$

with the initial condition

$$\lim_{t\downarrow 0} g^{(d)}_{k_1,k_2}(t,z_1,z_2] = 1.$$

2.4 Distribution of the number of end-nodes

The dynamics of the random tree evolution is controlled by the the end-nodes. It is evident that the living end-nodes are responsible for the development of a tree, while the dead end-nodes represent those points where the evolution was stopped. The probability distribution of the number of dead end-nodes $p_0^{(d)}(t,n)$ can be obtained by substitution $k=0$ into Eq. (5). It is easy to show that the generating function

$$g_0^{(d)}(t,z) = \sum_{n=0}^{\infty} p_0^{(d)}(t,n)\ z^n$$

satisfies the equation

$$g_0^{(d)}(t,z) = e^{-Qt} - (1-z)f_0(1-e^{-Qt}) + Q\ \int_0^t e^{-Q(t-t')}q\left[g_0^{(d)}(t',z)\right]\ dt'. \qquad (13)$$

In order to have an insight into the interplay between the living and dead end-nodes it seems to be useful to calculate the probability distribution of the random function

$$\mu(t,0) = \mu_\ell(t,0) + \mu_d(t,0)$$

and the joint distribution of $\mu_\ell(t,0)$ and $\mu_d(t,0)$.

With help of similar arguments used for the derivation of Eq. (7) it can be easily obtained the equation

$$g_0(t,z) = e^{-Qt}\ z - (1-z)f_0(1-e^{-Qt}) + Q\ \int_0^t e^{-Q(t-t')}q\left[g_0(t',z)\right]\ dt', \qquad (14)$$

for the generating function

$$g_0(t,z) = \sum_{n=0}^{\infty} \mathcal{P}\{\mu(t,0) = n|\mathcal{S}_0\}\ z^n = \sum_{n=0}^{\infty} p_0(t,n)\ z^n,$$

where $p_0(t,n)$ is the probability that the number of all end-nodes at $t \geq 0$ is equal to n provided that at $t=0$ the tree was in the state $\mathcal{S}_0$.

Now we would like to determine the joint distribution of $\mu_\ell(t,0)$ and $\mu_d(t,0)$, i.e. the probability

$$\mathcal{P}\{\mu_\ell(t) = n_1,\ \mu_d(t,0) = n_2|\mathcal{S}_0\} = p_0^{(\ell,d)}(t,n_1,n_2).$$

It is evident that

$$p_0^{(\ell,d)}(t,n_1,n_2) = e^{-Qt}\delta_{n_1,1}\delta_{n_2,0} +$$
$$+Q\int_0^t e^{-Q(t-t')}\ \left[f_0\delta_{n_1,0}\delta_{n_2,1} + \sum_{k=1}^{\infty} f_k\ R_0^{(\ell,d)}(t',n_1,n_2)\right]\ dt',$$

where

$$R_0^{(\ell,d)}(t',n_1,n_2) = \sum_{n_{11}+\cdots+n_{1k}=n_1}\ \sum_{n_{21}+\cdots+n_{2k}=n_2}\ \prod_{j=1}^{k} p_0^{(\ell,d)}(t',n_{1j},n_{2j}).$$

Simple calculations show that the generating function

$$g_0^{(\ell,d)}(t,z_1,z_2) = \sum_{n_1=0}^{\infty}\ \sum_{n_2=0}^{\infty} p_0^{(\ell,d)}(t,n_1,n_2)\ z_1^{n_1}\ z_2^{n_2}$$

satisfies the equation

$$g_0^{(\ell,d)}(t,z_1,z_2) = e^{-Qt}z_1 - f_0(1-z_2)(1-e^{-Qt}) +$$

$$+Q\int_0^t e^{-Q(t-t')}\ q\left[g_0^{(\ell,d)}(t,z_1,z_2)\right]\ dt', \tag{15}$$

which will be used in the next section for the determination of the correlation between the random variables $\mu_\ell(t,0)$ and $\mu_d(t,0)$.

Finally, we would like to cite the result of calculations concerning the probability to find the random function

$$\mu_d^{(o)}(t) = \sum_{k=1}^{\infty} \mu_d(t,k),$$

i.e. *the number of non end-nodes* at $t \geq 0$ to be equal to n, provided that at $t = 0$ the tree was in the state $\mathcal{S}_0$. Denote this probability by $p_d^{(o)}(t, n)$ and let us introduce the generating function

$$g_d^{(o)}(t, z) = \sum_{n=0}^{\infty} \mathcal{P}\{\mu_d^{(o)}(t) = n | \mathcal{S}_0\}\ z^n = \sum_{n=0}^{\infty} p_d^{(o)}(t, n)\ z^n.$$

It can be proven that it satisfies the integral equation

$$g_d^{(o)}(t, z) = e^{-Qt} + f_0(1-z)(1-e^{-Qt}) + zQ \int_0^t e^{-Q(t-t')}\ q \left[g_d^{(o)}(t', z)\right]\ dt', \quad (16)$$

which is equivalent to the differential equation

$$\frac{\partial g_d^{(o)}(t, z)}{\partial t} = Qf_0(1 - z) - g_d^{(o)}(t, z) + Qzq \left[g_d^{(o)}(t, z)\right]$$

with initial condition $\lim_{t \to 0} g_d^{(o)}(t, z) = 1$.

2.5 Average characteristics

It is difficult to find solutions of the generating function equations even in those cases when the distribution of ν is known and simple. [2] Therefore, in this section we would like to deal with the average properties of tree evolution, and will derive equations for the expectation values and variances of the number of nodes of different kind. In order to have an insight into the character of the stochastic interplay between the numbers of the living and dead end-nodes, we will investigate the time variation of the correlation between these nodes.

2.5.1 Dead nodes with k outgoing lines

Let the first task be the calculation of the time dependence of the expectation value and variance of dead nodes with $k \geq 0$ outgoing lines. By using the relation

$$\mathbf{E}\{\mu_d(t, k)\} = \left[\frac{\partial g_k^{(d)}(t, z)}{\partial z}\right]_{z=1} = m_1^{(d)}(t, k), \quad (17)$$

[2] As has been already mentioned the truncated ($\mathbf{t}$) arbitrary distribution of ν is one of the rare exceptions when the generating function equation can be exactly solved.

from the probability generating function (6) we obtain that

$$m_1^{(d)}(t,k) = f_k\ (1-e^{-Qt}) + q_1\ Q\ \int_0^t e^{-Q(t-t')}\ m_1^{(d)}(t',k)\ dt'. \tag{18}$$

The solution of this equation can be written in the form:

$$m_1^{(d)}(t,k) = \begin{cases} \frac{f_k}{1-q_1}\ \left[1-e^{-(1-q_1)Qt}\right], & \text{if } q_1 \neq 1, \\ \\ f_k\ Qt, & \text{if } q_1 = 1. \end{cases} \tag{19}$$

Since

$$\sum_{k=0}^{\infty} \mu_d(t,k) = \mu_d(t),$$

it is obvious that

$$\sum_{k=0}^{\infty} m_1^{(d)}(t,k) = m_1^{(d)}(t) = \begin{cases} \frac{1}{1-q_1}\ \left[1-e^{-(1-q_1)Qt}\right], & \text{if } q_1 \neq 1, \\ \\ Qt, & \text{if } q_1 = 1, \end{cases}$$

what has been already derived in [1]. By using this relation we can conclude that

$$\frac{m_1^{(d)}(t,k)}{m_1^{(d)}(t)} = f_k, \qquad \forall k \in \mathcal{Z}. \tag{20}$$

In order to calculate the variance $\mathbf{D}^2\{\mu_d(t,k)\}$ we need the second factorial moment $m_2^{(d)}(t,k)$. It can be shown that $m_2^{(d)}(t,k)$ satisfies the integral equation

$$m_2^{(d)}(t,k) = q_1\ Q\ \int_0^t e^{-Q(t-t')}\ m_2^{(d)}(t',k)\ dt' +$$

$$Q\ \int_0^t e^{-Q(t-t')}\ \left\{2k\ f_k\ m_1^{(d)}(t',k) + q_2\ \left[m_1^{(d)}(t',k)\right]^2\right\}\ dt'$$

the solution of which can be written in the form:

$$m_2^{(d)}(t,k) = \frac{f_k^2}{(1-q_1)^2}\left(2k + \frac{q_2}{1-q_1}\right)\ (1-e^{-\alpha t}) -$$

$$2\ \frac{f_k^2}{1-q_1}\left(2k + \frac{q_2}{1-q_1}\right)\ Qt\ e^{-\alpha t} + \frac{f_k^2}{(1-q_1)^3}\ q_2\ e^{-\alpha t}\ (1-e^{-\alpha t}), \tag{21}$$

for all $q_1 \neq 1$. Here the notation

$$\alpha = (1 - q_1)\, Q$$

has been used. If $q_1 = 1$, then the solution has the form:

$$m_2^{(d)}(t,k) = f_k^2\, (Qt)^2\, \left(k + \frac{1}{3}\, q_2\, Qt\right). \tag{22}$$

Taking into account that

$$\mathbf{D}^2\{\mu_d(t,k)\} = m_2^{(d)}(t,k) + m_1^{(d)}(t,k)\left[1 - m_1^{(d)}(t,k)\right],$$

where

$$m_1^{(d)}(t,k)\left[1 - m_1^{(d)}(t,k)\right] =$$

$$\frac{f_k}{1-q_1}\left(1 - \frac{f_k}{1-q_1}\right)\,(1 - e^{-\alpha t}) + \frac{f_k^2}{(1-q_1)^2}\, e^{-\alpha t}(1 - e^{-\alpha t}),$$

we obtain the following expression for the variance:

$$\mathbf{D}^2\{\mu_d(t,k)\} =$$

$$\frac{f_k}{1-q_1}\left[1 + \frac{f_k}{1-q_1}\left(2k - 1 + \frac{q_2}{1-q_1}\right)\right]\,(1 - e^{-\alpha t}) -$$

$$2\,\frac{f_k^2}{1-q_1}\left(2k + \frac{q_2}{1-q_1}\right)\, Qt\, e^{-\alpha t} +$$

$$\frac{f_k^2}{(1-q_1)^2}\left(1 + \frac{q_2}{1-q_1}\right)\, e^{-\alpha t}(1 - e^{-\alpha t}), \tag{23}$$

$$q_1 \neq 1 \quad \text{and} \quad \forall k \in \mathcal{Z},$$

where q_2 can be replaced by $\mathbf{D}^2\{\nu\} - q_1\,(1 - q_1)$. When the evolution is critical, i.e. when $q_1 = 1$, then we have

$$\mathbf{D}^2\{\mu_d(t,k)\} = f_k\, Qt\, \left[1 + (k-1)\, f_k\, Qt + \frac{1}{3}\, f_k\, \mathbf{D}^2\{\nu\}\, (Qt)^2\right], \tag{24}$$

$$\forall k \in \mathcal{Z}.$$

If the time t is converging to ∞ the Eqs. (23) and (24) show that the variance remains finite in the case of subcritical evolution only. Introducing the notation

$$1 + \frac{q_2}{1-q_1} = \frac{\mathbf{D}^2\{\nu\} + (1-q_1)^2}{1-q_1},$$

we obtain immediately the formula

$$\lim_{t\to\infty} \mathbf{D}^2\{\mu_d(t,k)\} =$$

$$\frac{f_k}{1-q_1}\left[1 + \frac{f_k}{1-q_1}\left(2(k-1) + \frac{q_2}{1-q_1}\right)\right], \qquad \forall k \in \mathcal{Z}, \tag{25}$$
$$\text{if} \quad q_1 < 1.$$

2.5.2 Properties of end-nodes

Now we would like to deal with some properties of end-nodes. The time dependence of the expectation of the number of end-nodes is one of the simplest characteristics of randomly evolving trees. It can be calculated from Eq. (14). By using the relation

$$\mathbf{E}\{\mu(t,0)\} = m_1(t,0) = \left[\frac{\partial g_0(t,z)}{\partial z}\right]_{z=1},$$

after some elementary mathematics we have

$$m_1(t,0) = \begin{cases} \frac{f_0}{1-q_1} + \frac{1-f_0-q_1}{1-q_1}\, e^{-(1-q_1)Qt}, & \text{if } q_1 \neq 1, \\ \\ 1 + f_0\, Qt, & \text{if } q_1 = 1. \end{cases} \tag{26}$$

It is interesting to note that the expectation value of the number of dead end-nodes $m_1^{(d)}(t,0)$ can be easily calculated from (26). By substituting $m_1^{(\ell)}(t,0) = e^{-(1-q_1)Qt}$ into the equation

$$m_1^{(d)}(t,0) = m_1(t,0) - m_1^{(\ell)}(t,0),$$

one gets the formula

$$m_1^{(d)}(t,0) = \begin{cases} \frac{f_0}{1-q_1}\left(1 - e^{-(1-q_1)Qt}\right), & \text{if } q_1 \neq 1, \\ \\ f_0\, Qt, & \text{if } q_1 = 1 \end{cases} \tag{27}$$

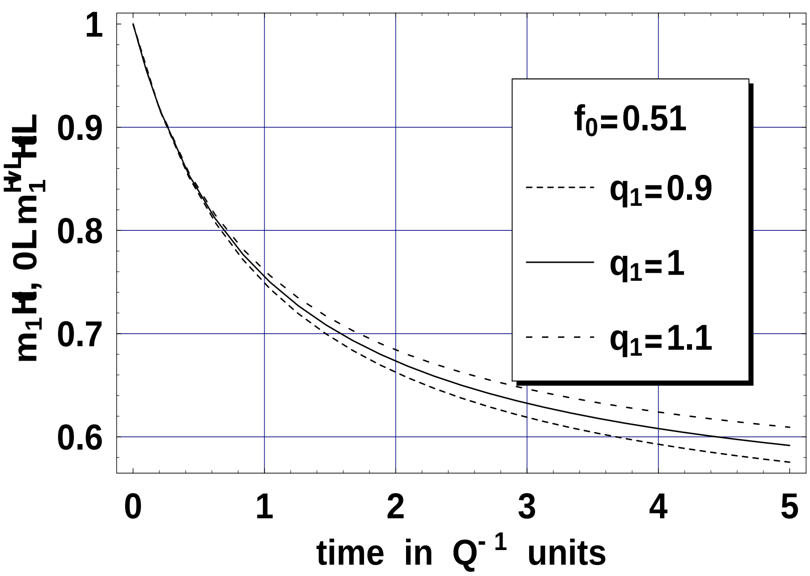


Figure 2: Time dependence of the ratio of the expected number of end-nodes $m_1(t,0)$ to that of all the nodes $m_1(t)$ in subcritical ($q_1 = 0.9$), critical ($q_1 = 1$) and supercritical ($q_1 = 1.1$) evolutions.

which is the same what follows from Eq. (19) when $k = 0$.

In many cases it seems to be important to know how the evolution process does alter the ratio of the expected number of end-nodes $m_1(t,0)$ to that of all the nodes $m_1(t)$. Taking into account that [3]

$$m_1(t) = \begin{cases} \frac{1-q_1\, e^{-(1-q_1)Qt}}{1-q_1}, & \text{if } q_1 \neq 1, \\ \\ 1 + Qt, & \text{if } q_1 = 1, \end{cases}$$

after elementary calculations we have

$$\frac{m_1(t,0)}{m_1(t)} = \begin{cases} \frac{f_0+(1-f_0-q_1)\, e^{-(1-q_1)Qt}}{1-q_1\, e^{-(1-q_1)Qt}}, & \text{if } q_1 \neq 1, \\ \\ \frac{1+f_0\, Qt}{1+Qt}, & \text{if } q_1 = 1, \end{cases}$$

[3]See Eq. (45) in [1]!

and Fig. 2 shows the ratio vs. time plots for values $q_1 = 0.9, 1, 1.1$ at $f_0 = 0.51$. As seen the curves are approaching the limit values

$$\lim_{t\to\infty} \frac{m_1(t,0)}{m_1(t)} = \begin{cases} f_0, & \text{if } q_1 \leq 1, \\ \\ 1 - \frac{1-f_0}{q_1}, & \text{if } q_1 > 1 \end{cases}$$

quite rapidly.[4]

In order to calculate the variance of the number of end-nodes we need the second factorial moment of $\mu(t,0)$. By using the relation

$$m_2(t,0) = \mathbf{E}\{\mu(t,0)\,[\mu(t,0)-1]\} = \left[\frac{\partial^2 g_0(t,z)}{\partial z^2}\right]_{z=1}$$

we obtain from Eq. (14) the integral equation

$$m_2(t,0) = q_1 Q \int_0^t e^{-Q(t-t')}\, m_2(t',0)\, dt' + q_2 Q \int_0^t e^{-Q(t-t')}\, [m_1(t',0)]^2\, dt',$$

the solution of that can be written in the form

$$m_2(t,0) = q_2 \left(\frac{f_0}{1-q_1}\right)^2 \frac{1-e^{-\alpha t}}{1-q_1} + 2q_2\, \frac{f_0(1-f_0-q_1)}{(1-q_1)^2}\, Qt\, e^{-\alpha t} +$$

$$+ q_2 \left(\frac{1-f_0-q_1}{1-q_1}\right)^2 e^{-\alpha t}\, \frac{1-e^{-\alpha t}}{1-q_1},$$

$$\text{if} \quad q_1 \neq 1.$$

When $q_1 = 1$, i.e. when the evolution is critical we get

$$m_2(t,0) = q_2 \left[Qt + f_0\,(Qt)^2 + \frac{1}{3} f_0^2\,(Qt)^3\right].$$

Finally we have the variance of $\mu(t,0)$ in the following form:

$$\mathbf{D}^2\{\mu(t,0)\} = \left[\frac{f_0(1-f_0-q_1)}{(1-q_1)^2} + \frac{q_2}{1-q_1}\left(\frac{f_0}{1-q_1}\right)^2\right](1-e^{-\alpha t}) +$$

[4]It is to note that $1 - \frac{1-f_0}{q_1} > f_0$ if $q_1 > 1$.

$$2q_2 \frac{f_0(1-f_0-q_1)}{(1-q_1)^2}\; Qt\; e^{-\alpha t} +$$

$$\left(1+\frac{q_2}{1-q_1}\right)\;\left(\frac{1-f_0-q_1}{1-q_1}\right)^2\; e^{-\alpha t}(1-e^{-\alpha t}), \tag{28}$$

$$\text{if} \quad q_1 \neq 1,$$

and

$$\mathbf{D}^2\{\mu(t,0)\} = (q_2 - f_0)\;(1 + f_0\; Qt)\; Qt + \frac{1}{3} q_2 f_0^2\; (Qt)^3, \tag{29}$$

$$\text{if} \quad q_1 = 1.$$

It can be easily proven that if $q_1 = 1$, then $q_2 - f_0 > 0$. Since $q_1 = \sum_{k=1}^{\infty} k f_k = 1$ and $\sum_{k=0}^{\infty} f_k = 1$ it is evident that

$$f_0 + f_1 + \sum_{k=2}^{\infty} f_k = f_1 + \sum_{k=2}^{\infty} k f_k,$$

and hence

$$f_0 = \sum_{k=2}^{\infty} (k-1)\; f_k.$$

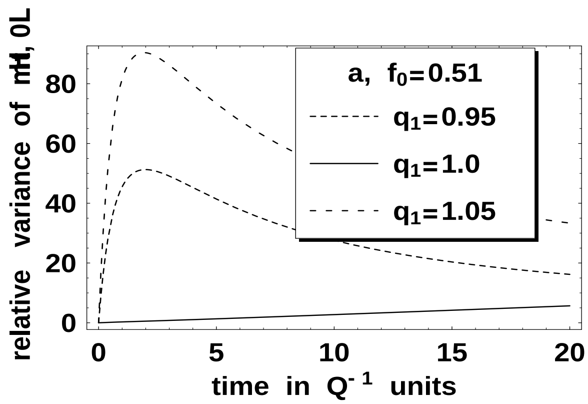
By using this expression of f_0 we find that

$$q_2 - f_0 = \sum_{k=2}^{\infty} [k(k-1) - (k-1)]\; f_k = \sum_{k=2}^{\infty} (k-1)^2\; f_k > 0. \qquad \text{Q.E.D.}$$

As follows from Eqs. (28) and (29) when $t \to \infty$ then the variance of $\mu(t,0)$ tends to finite limit value in the case of subcritical evolution only,

In order to show the main features of the time variation of fluctuations occurring in the end-node number of trees the relative variance and the relative dispersion of $\mu(t,0)$ vs. time have been calculated. The results are plotted in Fig. 3. It has been assumed that the distribution of ν is arbitrary in that sense as it was defined in section 2.1.1. In the upper part of the Fig. 3 one can see that the relative variance of $\mu(t,0)$ reaches a maximum just after the beginning of the process but in the case of subcritical and supercritical evolutions only. If $t \to \infty$, then we have

$$\lim_{t\to\infty} \frac{\mathbf{D}^2\{\mu(t,0)\}}{\mathbf{E}^2\{\mu(t,0)\}} = \begin{cases} \frac{1-f_0-q_1}{f_0} + \frac{q_2}{1-q_1}, & \text{if } q_1 < 1, \\ \\ \frac{q_2}{q_1-1} - 1, & \text{if } q_1 > 1, \end{cases}$$



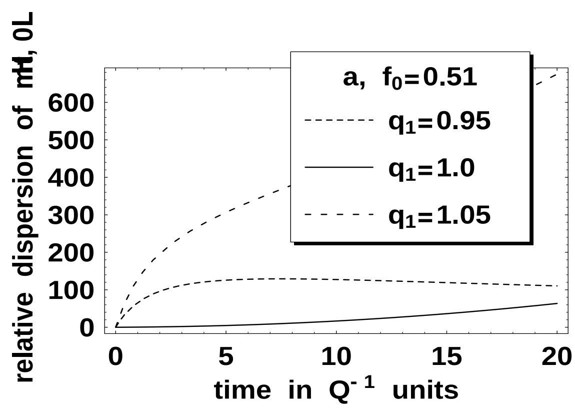


Figure 3: Relative variance (upper) and dispersion (lower) of the number of end-nodes vs. time in subcritical, critical and supercritical trees. The distribution of ν is arbitrary and $\mathbf{D}^2\{\nu\} = 0.9$.

The proof of the inequality $q_2 \geq q_1 - 1$ is very simple. If we substitute q_2 and q_1 by $\sum_{k=1}^{\infty} k(k-1)\, f_k$ and $\sum_{k=1}^{\infty} k\, f_k$, respectively, then we can write

$$\sum_{k=1}^{\infty} k(k-1)\, f_k - \sum_{k=1}^{\infty} k\, f_k + 1 \geq 0,$$

i.e.

$$\sum_{k=1}^{\infty} (k-1)^2\, f_k + f_0 \geq 0,$$

and this is trivially true.

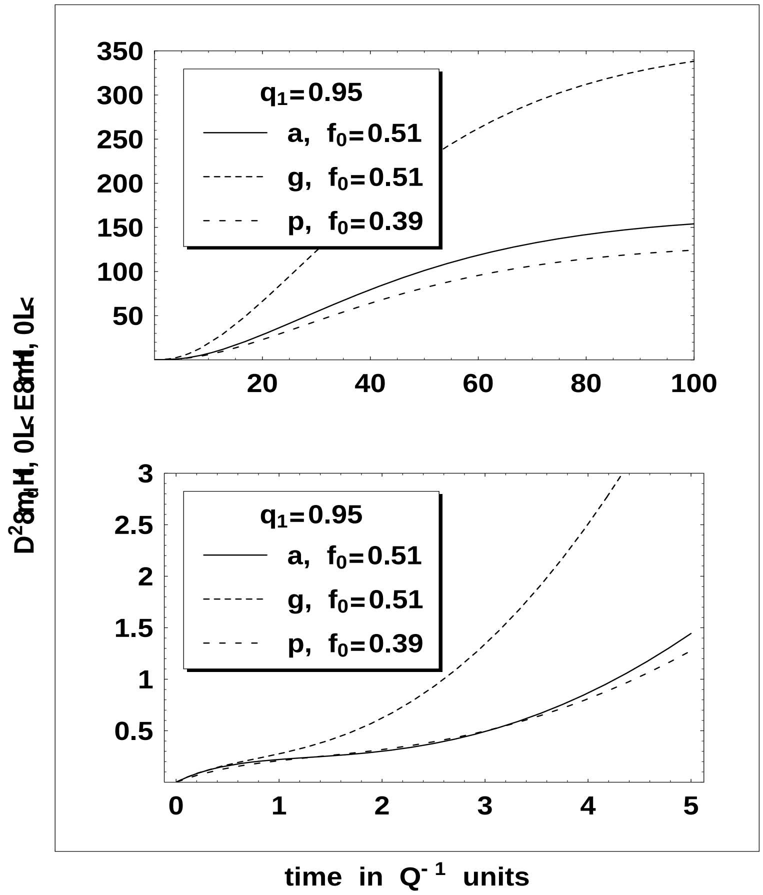


Figure 4: Relative dispersion of the number of end-nodes in subcritical trees vs. time. Curves **a, g**, and **p** correspond to the arbitrary, geometric and Poisson distributions of ν, respectively. In the case of arbitrary distribution: $\mathbf{D}^2\{\nu\} = 0.9$. Upper: $0 \leq Qt \leq 100$. Lower: $0 \leq Qt \leq 5$.

If the evolution is critical ($q_1 = 1$), then the relative variance

$$\frac{\mathbf{D}^2\{\mu(t,0)\}}{\mathbf{E}^2\{\mu(t,0)\}} = (q_2 - f_0)\,\frac{Qt}{1+f_0Qt} + \frac{1}{3}q_2 f_0^2\,\frac{(Qt)^3}{(1+f_0Qt)^2}$$

is increasing monotonously to infinity with t. The curves in the lower part of Fig. 3 show the time variation of the relative dispersion of $\mu(t,0)$ which

has a finite limit value

$$\lim_{t\to\infty} \frac{\mathbf{D}^2\{\mu(t,0)\}}{\mathbf{E}\{\mu(t,0)\}} = 1 + \frac{f_0}{1-q_1}\left(\frac{q_2}{1-q_1} - 1\right)$$

in the the subcritical evolution only.

The influence of the distribution law of ν on the relative dispersion can be seen in Fig. 4. The calculations have been carried out in the case of subcritical evolution. In the upper part of the figure it can be seen that each curves tends to a finite asymptotic value when $t \Rightarrow \infty$. It is remarkable that the geometric distribution of ν brings about much larger fluctuations in $\mu(t,0)$ than the arbitrary and Poisson distributions. The lower part of the figure shows the beginning of the time dependence which is reflecting the effect of two competing processes. One of them is the formation of new living nodes, while the other one is the death of end-nodes.

It seems to be worthwhile to calculate the time dependence of the expectation and variance of the number of dead end-nodes $\mu_d(t,0)$. Substituting $k=0$ into the Eqs. (19) and (23) it can be derived the relative variance

$$\frac{\mathbf{D}^2\{\mu_d(t,0)\}}{\mathbf{E}^2\{\mu_d(t,0)\}} = \left(\frac{q_2}{1-q_1} + \frac{1-q_1}{f_0} - 1\right)\,\frac{1}{1-e^{-\alpha t}} -$$

$$-2q_2\,Qt\,\frac{e^{-\alpha t}}{(1-e^{-\alpha t})^2} + \left(1 + \frac{q_2}{1-q_1}\right)\,\frac{e^{-\alpha t}}{1-e^{-\alpha t}},$$

$$\text{if} \quad q_1 \neq 1,$$

and

$$\frac{\mathbf{D}^2\{\mu_d(t,0)\}}{\mathbf{E}^2\{\mu_d(t,0)\}} = \frac{1}{3}q_2\,Qt - 1 + \frac{1}{f_0\,Qt},$$

$$\text{if} \quad q_1 = 1.$$

The relative variance of $\mu_d(t,0)$ converges to finite value, when $t \to \infty$, in both subcritical and supercritical evolutions but diverges if the evolution is critical. Let us introduce the notations

$$\lim_{t\to\infty} \frac{\mathbf{D}^2\{\mu_d(t,0)\}}{\mathbf{E}^2\{\mu_d(t,0)\}} = \begin{cases} rv_d^{(a)}, & \text{if } \nu \in \mathbf{a}, \\ \\ rv_d^{(g)}, & \text{if } \nu \in \mathbf{g}, \\ \\ rv_d^{(p)}, & \text{if } \nu \in \mathbf{p}, \end{cases}$$

and summarize the limit values of relative variances. We find that

$$rv_d^{(a)} = \begin{cases} \frac{1-q_1}{f_0} - 1 - q_1 + \frac{\mathbf{D}^2\{\nu\}}{1-q_1}, & \text{if } q_1 < 1, \\ \\ \infty, & \text{if } q_1 = 1, \\ \\ q_1 - 1 + \frac{\mathbf{D}^2\{\nu\}}{q_1-1}, & \text{if } q_1 > 1. \end{cases}$$

If the distribution of ν is geometric, then

$$rv_d^{(g)} = \begin{cases} q_1^2\,\frac{1+q_1}{1-q_1}, & \text{if } q_1 < 1, \\ \\ \infty, & \text{if } q_1 = 1, \\ \\ \frac{2q_1^2}{q_1-1} - 1, & \text{if } q_1 > 1, \end{cases}$$

while if it is of Poisson type, then

$$rv_d^{(p)} = \begin{cases} (1-q_1)\,e^{q_1} + \frac{q_1^2+q_1-1}{1-q_1}, & \text{if } q_1 < 1, \\ \\ \infty, & \text{if } q_1 = 1, \\ \\ \frac{q_1^2}{q_1-1} - 1, & \text{if } q_1 > 1. \end{cases}$$

Finally, it seems to be interesting to look at the time dependence of the relative dispersion the number of dead end-nodes. It follows from Eqs. (19) and (23) that

$$\frac{\mathbf{D}^2\{\mu_d(t,0)\}}{\mathbf{E}\{\mu_d(t,0)\}} = 1 + \frac{f_0}{1-q_1}\,\left(\frac{q_2}{1-q_1} - 1\right) +$$

$$+\frac{f_0}{1-q_1}\,\left(\frac{q_2}{1-q_1} + 1\right)\,e^{-\alpha t} - 2f_0\,\frac{q_2}{1-q_1}\,Qt\,\frac{e^{-\alpha t}}{1-e^{-\alpha t}},$$

$$\text{if} \quad q_1 \neq 1,$$

and

$$\frac{\mathbf{D}^2\{\mu_d(t,0)\}}{\mathbf{E}\{\mu_d(t,0)\}} = 1 + f_0\,Qt\left(\frac{1}{3}q_2\,Qt - 1\right),$$

$$\text{if} \quad q_1 = 1.$$

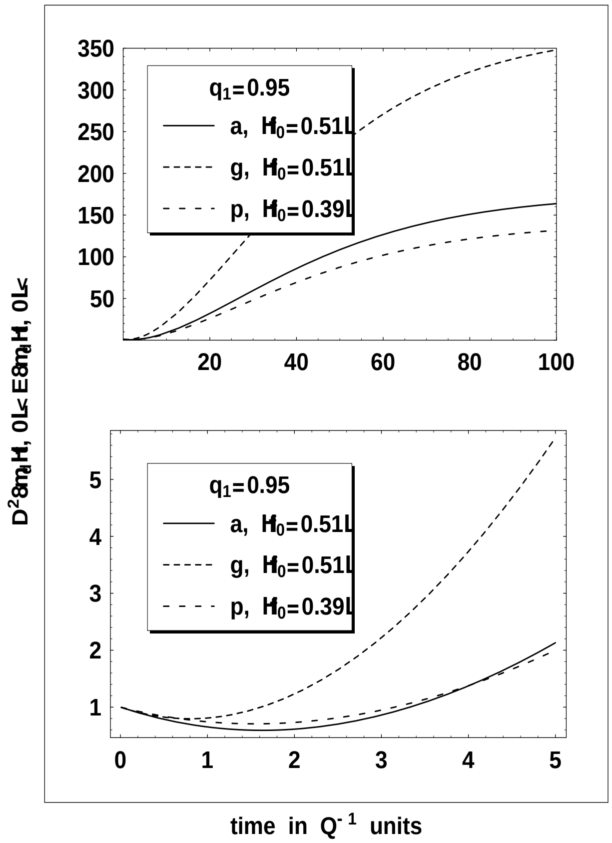


Figure 5: Relative dispersion of the number of dead end-nodes vs. time in subcritical trees. Curves **a, g**, and **p** correspond to the arbitrary, geometric and Poisson distributions of ν, respectively. In the case of arbitrary distribution: $\mathbf{D}^2\{\nu\} = 0.9$ and $f_0 = 0.51$. Upper: $0 \leq Qt \leq 100$. Lower: $0 \leq Qt \leq 5$.

Fig. 5 shows the time dependence of the relative dispersion of the dead end-nodes $\mu_d(t, 0)$ in subcritical evolution for all the three (**a, g, p**) distributions of ν. In the lower part of the figure one can see that the relative dispersion curves have minimum just after the beginning of the process.(In the case of the arbitrary distribution of ν the minimum sites are given in Table I.)

Table I: Sites of the minimum in relative dispersion vs. time curves at several q_1 and $\mathbf{D}^2\{\nu\}$ values. (Time is measured in Q^{-1} units and $\nu \in \mathbf{a}$.)

$\mathbf{D}^2\{\nu\}\backslash q_1$	0.90	0.95	1.00	1.05	1.10
0.90	1.7724	1.7223	1.6617	1.6069	1.5441
1.00	1.5849	1.5448	1.5000	1.4514	1.3999
1.10	1.4332	1.4006	1.3636	1.3233	1.2803

3 Lifetime of trees

3.1 General considerations

It is obvious that the evolution of a random tree will stop at that time instant θ which satisfies with probability one the equation $\mu_\ell(\theta) = 0$. The random variable θ is called *lifetime of the tree*. In order to determine its distribution function

$$\mathcal{P}\{\theta \leq t|\mathcal{S}_0\} = L(t), \tag{30}$$

one has to recognize that the probability

$$\mathcal{P}\{\mu_\ell(t) = 0|\mathcal{S}_0\} = p^{(\ell)}(t, 0)$$

to find zero living node at time moment $t \geq 0$ in a tree is the same as the probability that the lifetime θ of that tree is not larger than $t \geq 0$, therefore, one can write

$$\mathcal{P}\{\theta \leq t|\mathcal{S}_0\} = \mathcal{P}\{\mu_\ell(t) = 0|\mathcal{S}_0\},$$

i.e.,

$$L(t) = p^{(\ell)}(t, 0) = \lim_{z\downarrow 0} \; g^{(\ell)}(t, z). \tag{31}$$

It is clear that if $0 < t_1 \leq t_2$ then $L(t_1) \leq L(t_2)$, i.e., $L(t)$ is a non-decreasing function of its argument, hence the limit relation

$$\max_{0<t\leq\infty} \; L(t) = \lim_{t\to\infty} \; L(t) = L_\infty \leq 1 \tag{32}$$

must be valid. We will call the quantity L_∞ *dying-out-probability*, and prove the following theorem:

Theorem 1 *If $q_1 \leq 1$, i.e., the random evolution is not supercritical, then $L_\infty = 1$, while if $q_1 > 1$, i.e., the evolution is supercritical, then L_∞ is equal to the single, smaller than 1, non-negative root of the function $\psi(y) = q(y) - y, \;\; y \in [0, 1]$.* [5]

[5] It is evident that $\psi(1) = 0$.

Proof. For the proof we exploit the fundamental property of the generating function $g^{(\ell)}(t,z)$ which is expressed by the equation

$$g^{(\ell)}(t+u,z) = g^{(\ell)}\left[t, g^{(\ell)}(u,z)\right].$$

Applying the relation (31) we have

$$L(t+u) = g^{(\ell)}\left[t, L(u)\right]$$

and since

$$\lim_{u\to\infty} L(u) = \lim_{u\to\infty} L(t+u) = L_\infty,$$

we can write for every $t \geq 0$ that

$$L_\infty = g^{(\ell)}(t, L_\infty).$$

Putting $g^{(\ell)}(t, L_\infty)$ into

$$\frac{dg^{(\ell)}(t,z)}{dt} = Q\ q\left[g^{(\ell)}(t,z)\right] - Q\ g^{(\ell)}(t,z) \tag{33}$$

derived from (2), we obtain

$$q\left(L_\infty\right) - L_\infty = 0. \tag{34}$$

Considering that $q(y)$ is probability generating function, i.e., $\lim_{y\uparrow 1}\ q(y) = 1$, according to a well-known theorem of generating functions [3], it is clear that if $q_1 > 1$ then Eq. (34) — besides the trivial fixed-point 1 — must have an other, smaller than 1, non-negative fixed-point L_∞ too, and this is what we wanted to prove. Q.E.D. [6]

Let us introduce the probability

$$S(x) = 1 - L(x), \tag{35}$$

to find the tree at the time instant $x = Qt$ in living state. The $S(x)$ will be called *survival probability.* By taking into account the properties of $L(x)$ one obtains

$$\lim_{Qt\to\infty} S(x) = \begin{cases} 0, & \text{if } q_1 \leq 1, \\ S_\infty = 1 - L_\infty, & \text{if } q_1 > 1. \end{cases} \tag{36}$$

[6]The proof of this and the following theorems is based on Sjewastjanow's ideas [4].

By using Eqs. (33), (31), and (35) one gets

$$\frac{dS}{dx} = -q(1-S) - S + 1, \tag{37}$$

which has the solution

$$x = \int_{S(x)}^{1} \frac{dy}{q(1-y) + y - 1}, \tag{38}$$

if the initial condition is $S(0) = 1$. Now, we would like to derive *asymptotic expressions* of $S(x)$ for large $x = Qt$ in the cases of subcritical, critical and supercritical evolution.

Theorem 2 *If the integral*

$$\int_0^1 \frac{q(1-y) + q_1 y - 1}{y[q(1-y) + y - 1]}\, dy = -\log K \tag{39}$$

is finite, then in the case of subcritical evolution the survival probability $S(x)$ has the following form:

$$S(x) = K\, e^{-(1-q_1)x}\, [1 + o(1)] \tag{40}$$

when $x = Qt \Rightarrow \infty$.

Proof. The proof is simple: the identity

$$\log \frac{e^{-(1-q_1)x}}{S(x)} = (q_1 - 1)\, x - \log S(x),$$

by using the Eq. (38), can be rewritten in the form

$$\log \frac{e^{-(1-q_1)x}}{S(x)} = (q_1 - 1) \int_{S(x)}^{1} \frac{dy}{q(1-y) + y - 1} + \int_{S(x)}^{1} \frac{dy}{y} =$$

$$\int_{S(x)}^{1} \frac{q(y-1) + q_1\, y - 1}{y\, [q(1-y) + y - 1]}\, dy = k(x),$$

and hence

$$S(x) = e^{-k(x)}\, e^{-(1-q_1)x}.$$

From this we find
$$\lim_{x\to\infty} \frac{S(x)}{e^{-(1-q_1)x}} = e^{-k(\infty)},$$
where
$$k(\infty) = \int_0^1 \frac{q(1-y)+q_1 y - 1}{y[q(1-y)+y-1]}\, dy = -\log K$$
since $S(\infty) = 0$ if $q_1 < 1$. According to the assumption (39) we can write
$$S(x) = e^{-k(\infty)}\, e^{-(1-q_1)x}\, [1+o(1)] = K\, e^{-(1-q_1)x}\, [1+o(1)],$$
and this is what we wanted to prove. Q.E.D.

It is to mention that in the case of **t** distribution of ν from Eq. (39) we obtain
$$K = \frac{1}{1+\frac{1}{2}\,\frac{q_2}{1-q_1}},$$
and so
$$S(x) = \frac{e^{-(1-q_1)x}}{1+\frac{1}{2}\,\frac{q_2}{1-q_1}}\, [1+o(1)].$$

Theorem 3 *If $q_2 < \infty$ and $q_1 = 1$, then the asymptotic expression for the survival probability is given by*
$$S(x) = \frac{2}{q_2\, x}\, [1+o(1)] \tag{41}$$
when $x = Qt \Rightarrow \infty$.

Proof. By using the series expansion theorem according to which
$$q(1-S) = 1 - S + \frac{1}{2}\, q''[b(x)]\, S^2,$$
where $1 - S \leq b(x) < 1$, from Eq. (37) we obtain
$$\frac{dS}{dx} = -\frac{1}{2}\, q''[b(x)]\, S^2.$$
Since if $x \Rightarrow \infty$, then $S(x) \Rightarrow 0$ we can write
$$q''[b(x)] = q_2 + \epsilon(x), \qquad \text{where} \qquad \lim_{x\to\infty} \epsilon(x) = 0.$$

Taking into account this expression for $q''[b(x)]$ we have

$$\frac{dS}{dx} = -\frac{1}{2}\, q_2\, S^2 - \frac{1}{2}\, \epsilon(x)\, S^2,$$

the solution of which can be written in the form

$$S(x) = \left[\frac{1}{2}\, q_2\, x + \frac{1}{2}\, \int_0^x \epsilon(v)\, dv + C\right]^{-1}.$$

The initial condition $S(0) = 1$ results in $C = 1$, and therefore

$$S(x) = \frac{2}{q_2\, x}\, \left[1 + \frac{2}{q_2 x} + \frac{1}{q_2 x}\, \int_0^x \epsilon(v)\, dv\right]^{-1}.$$

By applying the L'Hospital rule we find

$$\lim_{x \to \infty} \frac{1}{x}\, \int_0^x \epsilon(v)\, dv = \lim_{x \to \infty} \epsilon(x) = 0,$$

and hence

$$S(x) = \frac{2}{q_2\, x}\, [1 + o(1)],$$

what we wanted to prove. Q.E.D.

The third task is to derive the asymptotic expression for $S(x)$ in the case of supercritical evolution.

Theorem 4 *If the integral*

$$\int\limits_{S_\infty}^{1} \frac{q(1-y) + y - 1 - (q_1 - 1)\, (y - S_\infty)}{(y - S_\infty)\, [q(1-y) + y - 1]}\, dy = r(\infty) \tag{42}$$

is finite in the case of $q_1 > 1$, *then*

$$S(x) = S_\infty + (1 - S_\infty)\, e^{-r(\infty)}\, e^{-(q_1 - 1)x}\, [1 + o(1)], \tag{43}$$

for $x = Qt \Rightarrow \infty$ *where* $S_\infty = 1 - L_\infty < 1$ *is the limit value of the survival probability.*

Proof. From the identity

$$\log \frac{e^{-(q_1-1)x}}{S(x)-S_\infty} = -(q_1-1)\, x - \log[S(x)-S_\infty]$$

we obtain

$$\log \frac{e^{-(q_1-1)x}}{S(x)-S_\infty} = -(q_1-1) \int\limits_{S(x)}^{1} \frac{dy}{q(1-y)+y-1} + \int\limits_{S(x)-S_\infty}^{1} \frac{dy}{y}, \tag{44}$$

and since

$$\int\limits_{S(x)-S_\infty}^{1} \frac{dy}{y} = \int\limits_{S(x)}^{1} \frac{dy}{y-S_\infty} + \log \frac{1}{S(x)-S_\infty},$$

we can rewrite Eq. (44) in the form

$$\log \frac{1-S_\infty}{S(x)-S_\infty}\, e^{-(q_1-1)x} = r(x),$$

where

$$r(x) = \int\limits_{S(x)}^{1} \frac{q(1-y)+y-1-(q_1-1)\,(y-S_\infty)}{(y-S_\infty)\,[q(1-y)+y-1]}\, dy,$$

and so we have

$$S(x) = S_\infty + (1-S_\infty)\, e^{-r(x)}\, e^{-(q_1-1)x}.$$

It has been assumed $r(\infty)$ to be finite and so

$$\lim_{x\to\infty} \frac{S(x)-S_\infty}{e^{-(q_1-1)x}} = (1-S_\infty)\, e^{-r(\infty)} < \infty.$$

Thus the theorem is proved. Q.E.D.

In the case of **t** distribution of ν we find immediately that

$$S_\infty = 2\frac{q_1-1}{q_2}, \qquad \text{and} \qquad r(\infty) = -\log S_\infty,$$

and hence obtain

$$S(x) = 2\frac{q_1-1}{q_2} + 2\frac{q_1-1}{q_2}\left(1-2\frac{q_1-1}{q_2}\right) e^{-(q_1-1)x}\,[1+o(1)],$$

when $x = Qt \Rightarrow \infty$.

3.2 Some exactly solvable models

In order to demonstrate the characteristic features of the lifetime of random trees we will use such probabilities $f_k, \quad k = 0, 1, \ldots$ for the offspring numbers ν that make possible to solve exactly the equation (37) of the survival probability. In the following, two special cases will be investigated, namely

$$q(z) = \begin{cases} 1 + q_1(z-1) + \frac{1}{2}\, q_2\, (z-1)^2, & \text{model } \mathbf{t}, \\ \\ [1 + q_1(1-z)]^{-1}, & \text{model } \mathbf{g}. \end{cases} \tag{45}$$

It is clear that the first case corresponds to the zero-one-two, while the second to the geometric distribution of the offspring number ν.

3.2.1 Survival probabilities

By using Eq. (37) in the model $\mathbf{t}$ we obtain

$$\frac{dS}{dx} = -(1-q_1)\, S - \frac{1}{2}\, q_2\, S^2, \tag{46}$$

and taking into account the initial condition $S(0) = 1$ we have the solution in the form

$$S(x) = \begin{cases} e^{-(1-q_1)x} \left[1 + \frac{q_2}{2(1-q_1)}\, (1 - e^{-(1-q_1)x})\right]^{-1}, & \text{if } q_1 < 1, \\ \\ \frac{2}{2 + q_2\, x}, & \text{if } q_1 = 1, \\ \\ 2\, \frac{q_1 - 1}{q_2}\, \left[1 + (1 - 2\, \frac{q_1-1}{q_2})\, e^{-(q_1-1)x}\right]^{-1}, & \text{if } q_1 > 1. \end{cases} \tag{47}$$

The corresponding equation in the model $\mathbf{g}$ (i.e. when the distribution of ν is geometric) is given by

$$\frac{dS}{dx} = -\frac{1}{1 + q_1 S} - S + 1 = -S\, \frac{1 - q_1(1-S)}{1 + q_1 S}, \tag{48}$$

the solution of which can be easily obtained in inverse form

$$x(S) = \begin{cases} \frac{1}{1-q_1}\, \log\frac{1}{S} + \frac{q_1}{1-q_1}\, \log[1 - q_1(1-S)], & \text{if } q_1 \neq 1, \\ \\ \frac{1-S}{S} + \log\frac{1}{S}, & \text{if } q_1 = 1, \end{cases} \tag{49}$$

satisfying the initial condition $S(0) = 1$.

The density function of the lifetime measured in Q^{-1} units can be calculated by using the relation

$$\ell(x) = \frac{dL(x)}{dx} = -\frac{dS(x)}{dx}.$$

It is elementary to show that the density function is decreasing monotonously from $\ell(0) = f_0$ to zero. In Fig. 6 one can see the density function curves versus time $x = Qt$ for subcritical, critical and supercritical trees and for both distributions of ν.

Let us calculate now the characteristic function of the random variable $\vartheta = Q\theta$, i.e. of the tree lifetime measured in Q^{-1} units. Since the moments $\mathbf{E}\{\vartheta^j\},\ \ j = 1, 2, \ldots$ do not exist if the $q_1 \geq 1$, the calculations are restricted to the case when $q_1 < 1$. One can write

$$\varphi(\omega) = \mathbf{E}\{e^{-\omega\vartheta}\} = \int_0^\infty e^{-\omega x}\ dL(x) = -\int_0^\infty e^{-\omega x}\ dS(x) = \int_0^1 e^{-\omega x(y)}\ dy,$$

where ω is a complex number with $\Re\omega \geq 0$. Performing the substitution

$$x(y) = -\frac{1}{1-q_1}\ \log y\ \frac{1 + 2\frac{1-q_1}{q_2}}{y + 2\frac{1-q_1}{q_2}}$$

in the model $\mathbf{t}$ one obtains the characteristic function

$$\varphi_t(\omega) = (1+\gamma)^{\omega\beta}\ \int_0^1 \left[\frac{y}{y+\gamma}\right]^{\omega\beta}\ dy, \tag{50}$$

where

$$\beta = (1-q_1)^{-1} \qquad \text{and} \qquad \gamma = 2\frac{1-q_1}{q_2}, \qquad q_1 < 1.$$

In the model $\mathbf{g}$ we should substitute for x the expression given by (49), i.e.

$$x(y) = \frac{1}{1-q_1}\ \log\frac{1}{y} + \frac{q_1}{1-q_1}\ \log[1 - q_1(1-y)],$$

and we find

$$\varphi_g(\omega) = \int_0^1 (1-y)^{\omega\beta}\ (1-q_1 y)^{\omega(1-\beta)}\ dy, \qquad q_1 < 1. \tag{51}$$

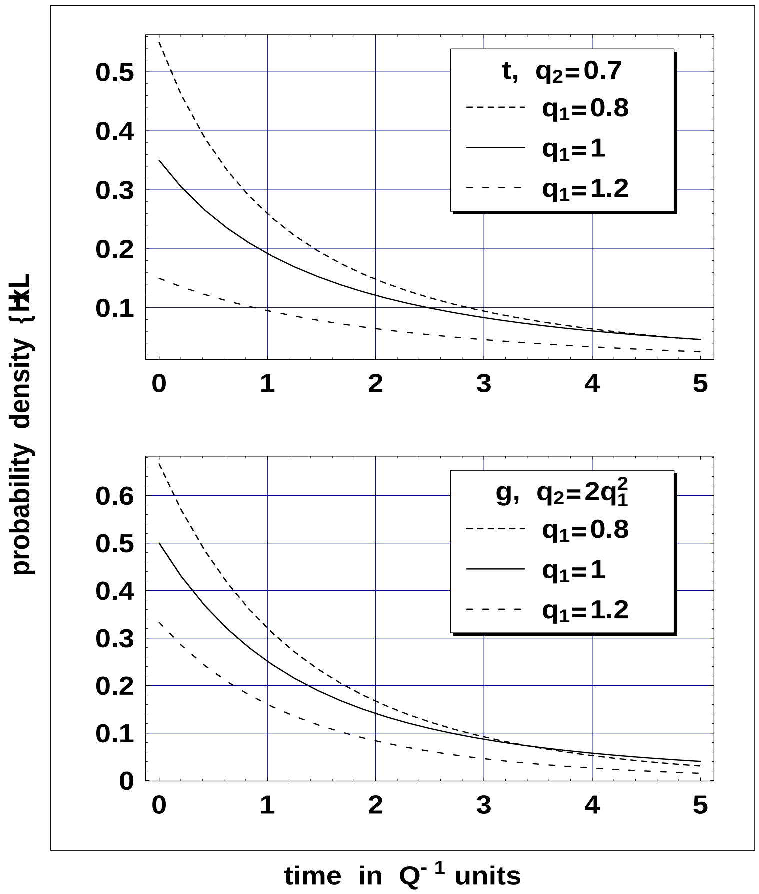


Figure 6: Time dependence of the density function $\ell(x)$ of the tree lifetime ϑ in models **t** and **g**.

3.2.2 Expectation and variance of the lifetime

From the characteristic functions $\varphi_t(\omega)$ and $\varphi_g(\omega)$ it can be easily calculated both the expectation value and the variance of the lifetime ϑ of a subcritical tree. In the model **t** for the expectation value one has

$$\mathbf{E}\{\vartheta_t\} = -\left(\frac{d\varphi_t(\omega)}{d\omega}\right)_{\omega=0} = \beta \int_0^1 \log\frac{y+\gamma}{y\,(1+\gamma)}\,dy =$$

$$\beta\,\gamma\,\log\left(1+\frac{1}{\gamma}\right) = \frac{2}{q_2}\,\log\left(1+\frac{1}{2}\,\frac{q_2}{1-q_1}\right). \tag{52}$$

while in the model **g** one finds expression

$$\mathbf{E}\{\vartheta_g\} = -\left(\frac{d\varphi_g(\omega)}{d\omega}\right)_{\omega=0} = 1 - \log(1 - q_1). \tag{53}$$

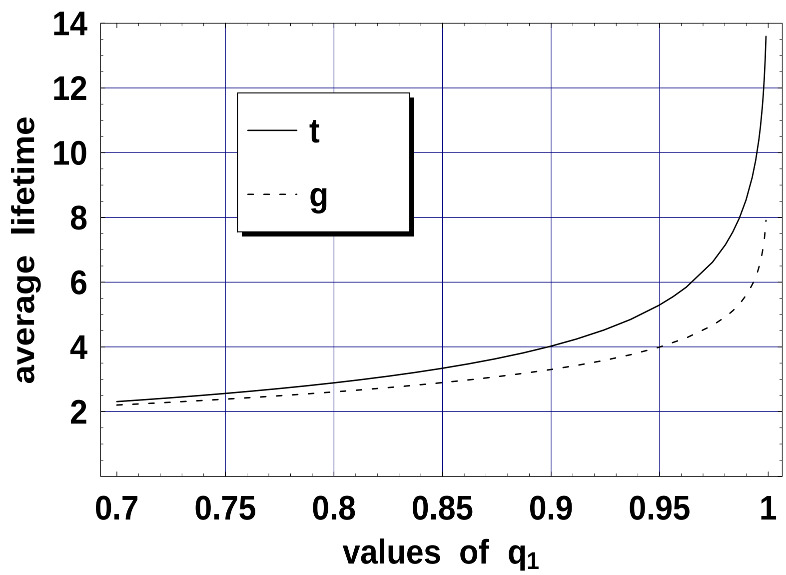


Figure 7: Dependence of the average value of the tree lifetime ϑ on $q_1 < 1$ in models **t** and **g**.

The curves in Fig. 7 show the dependence of the expectation of the tree lifetime on the average branching parameter q_1 in both models **t** and **g**. One can observe the difference between the two curves to be unimportant. In both cases if $q_1 \Rightarrow 1$ then the expectation value becomes infinite like $\log(1-q_1)^{-1}$.

For the calculation of the variance of the tree lifetime we need the second moment of ϑ which can be immediately obtained from the characteristic function. In the case of the model **t** we have

$$\mathbf{E}\{\vartheta_t^2\} = \left[\frac{d^2\varphi_t(\omega)}{d\omega^2}\right]_{\omega=0} = \beta^2 \int_0^1 \left[\log \frac{y+\gamma}{y\,(1+\gamma)}\right]^2 dy.$$

By using some well known integral relations it can be shown that

$$\mathbf{E}\{\vartheta_t^2\} = -\beta^2\,\gamma \left\{\left[\log\left(1+\frac{1}{\gamma}\right)\right]^2 + 2\,Li_2\left(-\frac{1}{\gamma}\right)\right\},$$

where $Li_2(u) = \sum_{k=1}^{\infty} u^k/k^2$ is the so called Jonquière's (dilogarithm)

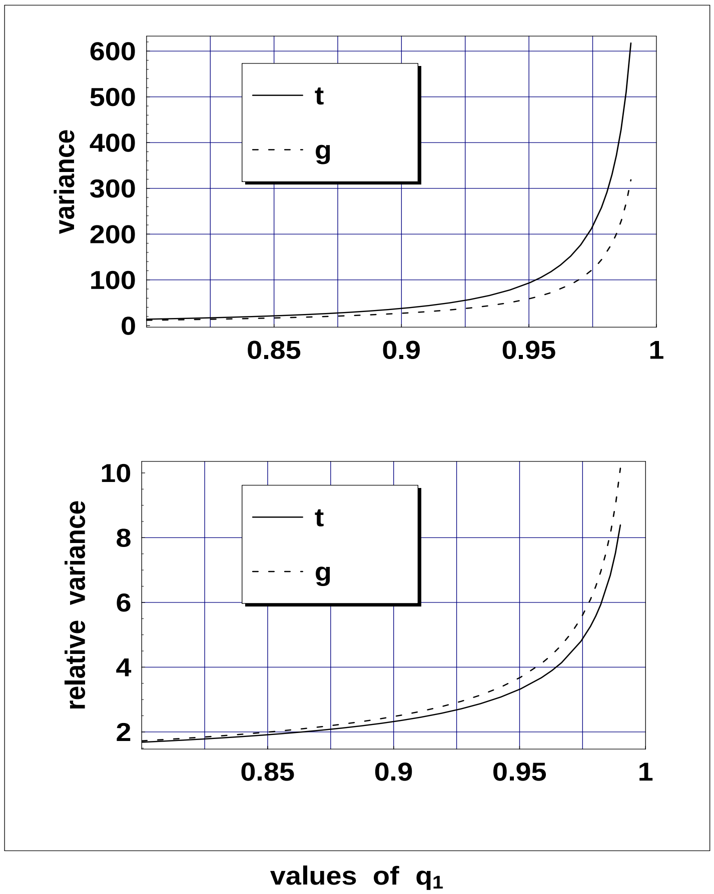


Figure 8: Dependence of the variance and the relative variance of the tree lifetime ϑ on $q_1 < 1$ in models **t** and **g**.

function. Taking into account this form of the second moment we can write

$$\mathbf{D}^2\{\vartheta_t\} = -\left(\frac{2}{q_2}\right)^2 \left(1 + \frac{1}{2}\,\frac{q_2}{1-q_1}\right) \left[\log\left(1 + \frac{1}{2}\,\frac{q_2}{1-q_1}\right)\right]^2 -$$

$$\frac{4}{q_2\,(1-q_1)}\, Li_2\left(-\frac{1}{2}\frac{q_2}{1-q_1}\right). \tag{54}$$

Performing similar calculations in the case of the model $\mathbf{g}$ for the variance of ϑ_g we obtain

$$\mathbf{D}^2\{\vartheta_g\} = 1 - \frac{1}{1-q_1}\ [\log(1-q_1)]^2 - \frac{2}{1-q_1}\ Li_2\left(-\frac{q_1}{1-q_1}\right). \tag{55}$$

In Fig. 8 one can see the dependence of the variance as well as the relative variance of the tree lifetime on the parameter q_1 in the cases of both models $\mathbf{t}$ and $\mathbf{g}$. When q_1 is approaching to 1 from below the fluctuation of the tree lifetime becomes unlimitedly large, and so in the vicinity of the critical state the average lifetime loses almost completely its information content.

4 Concluding remarks

The probability distribution of the number of nodes with $k \geq 0$ outgoing lines has been investigated in randomly evolving trees defined in [1]. Special attention was paid on the stochastic properties of end-nodes. We found that the birth and death of end-nodes in randomly evolving trees are playing a decisive role in the dynamics of the process.

It is remarkable that the *relative variance* of the number of end-nodes vs. time has well-defined maximum when the evolution is either subcritical or supercritical. In the case of critical evolution the relative variance increases monotonously with the time. On the contrary, the *relative dispersion* of the number of dead end-nodes vs. time has a minimum just after the beginning of the evolution. The minimum can be seen in each of evolution states.

We defined the lifetime of randomly evolving trees and derived a non-linear differential equation for the probability that the lifetime is larger than a given positive real number x. Three theorems have been proven to obtain asymptotic expressions for the survival probability of subcritical, critical and supercritical trees. Though the average lifetime of supercritical trees is always infinite it has been shown that the probability of finite lifetime of supercritical trees is larger than zero. In other words, randomly evolving supercritical trees may have finite lifetime.

References

[1] L. Pál, Randomly evolving trees I, arXiv:cond-mat/0205650

[2] L. Pál, *The Foundation of the Probability Calculus and Statistics* (Akadémiai Kiadó, Budapest, 1995)

[3] L. Pál, *Properties of Generating Functions* (Internal Report, Budapest, 1994)

[4] B. Sewastjanow, *Verzweigungsprozesse* (Akademie-Verlag, Berlin, 1974)